

学科排名: 利大还是弊大

——对我国学科评估特征、正当性与机理的省思

王洪才

(厦门大学 高等教育发展研究中心 福建 厦门 361005)

摘 要: 学科评估在我国高教界受到了异乎寻常的重视,因为它与“双一流”建设的紧密相关。通过对学科评估目的、学科排名手段和学科趋同化结果的考察可以发现,我国学科评估运作是行政主导式的,具有集中性统一化特点,它通过学科排名方式来实践。学科排名虽然极大地调动了高校参与的热忱,但却造成了盲目攀比和趋同化效应。尽管第四轮学科评估在技术上做了重大改进,但根本特征未变,且所采用的“绑定原则”与“自愿参评”初衷相悖,坚持“一级学科”评估原则也不利于学科走内涵式发展道路。高校学科评估应当注重诊断而非排名,高校自主权应当受到重视,大学个性和学科特色应受到高度尊重。

关键词: 学科评估; 学科排名 “双一流”建设

中图分类号: G640 **文献标识码:** A **文章编号:** 0438-0460(2019) 01-0078-09

学科建设可谓是当前我国高校工作的重心所在,这显然与我国正在大力推进的“双一流”建设背景直接相关,当然也与人们越来越清晰地认识到“学科作为大学的基本单元”的意义有关。人们发现,如果学科建设得不好,大学就不可能建设好;而且只有个别学科建设得好也不足以代表大学整体建设得好。这告诉人们,学科建设与大学发展之间既是统一又是相互区别的,这是“双一流”建设的基本逻辑。^[1] 我们知道,一流大学是一系列高水平学科的集合,没有学科群体的成功,就不可能有真正的一流大学。^[2] 大学成功的关键显然在于管理水平,如果没有良好的大学文化氛围,没有良性的学术机制,一流学科就不可能诞生,当然也不可能产生一流大学。如果说个别学科的成功存在偶然性的话,那么学科群体的成功就具有必然性了,因为它是制度环境的结果,是大学治理水平的反映,也是大学文化积淀的折射。有鉴于此,我国采取大学建设与学科建设并举策略来推进高等教育强国建设,这既是对大学发展一般规律的反映,也是学科发展特殊规律的展示。毋庸置疑,大学建设必须尊重学科个性,不能搞简单化的、一刀切模式。只有充分尊重学科个性,才能发挥学科自身的创造潜能,才能展现学科之间的互补性,才能发挥其相互促进的作用。为此,一流大学建设呼唤大学治理水平的现代化。如果没有宽松制度环境,学科建设就不可能成功,大学自身更不可能成功,因为大学发展是以尊重学科多样性为前提的,只有百家争鸣,才可能出现学术繁荣。这也说明,无论什么样的大学制度设计,都必须以学术繁荣为旨归,以促进学科发展为出发点。大学学科评估制度自然也不例外,也必须以尊重学科多样性特征为前提,以促进大学分类有序发展为旨归,不然,学科评估的价值就值得质疑。

收稿日期: 2018-09-27

基金项目: 教育部人文社科重点研究基地重大项目“中国特色的大学内部治理结构与质量保障机制建设研究”(18JJD880005)

作者简介: 王洪才,男,河北永年人,厦门大学教育研究院教授、博士生导师。

一、我国学科评估模式的特征及其缺陷

学科发展要不要评估?这是一个毋庸置疑的问题。学科建设总有自己的目标,它是否达到了计划目标要求,是否需要调整,都需要进行一个客观评估,如果没有科学的评估就难以知道学科建设的具体进展,甚至可能会迷失学科建设的方向。因此,人们对是否需要学科评估并无异议。^[3]但该如何进行评估?怎样的评估才是科学的公正的?对此就有很大的分歧。^[4]人们都希望进行科学公正的评估,都希望获得公平的对待,并希望评估对自身学科建设起到推动作用而不是伤害,因此,采用什么样的评估机制就成为一个核心问题。

我国目前所采用的是一种集中性统一化的学科评估模式,它虽然反映了我国高校管理体制根本特征,但并非是最有利于学科发展的机制。所谓集中性统一化评估,是指由一个官方或半官方的权威部门进行组织,针对所有学科,在集中的时间内,实行统一的评估。在评估中,针对不同学科采用不同的指标体系,从而在全国范围内形成一个高度集中的、分学科的、固定期限内实施的学科评估。它的优势非常明显:1. 指标统一,一视同仁,表明了评估规则客观公正;2. 时间集中,效率高,从而便于学校统一布置安排;3. 结果公布,指标透明,从而便于学校之间相互监督;4. 信息采集全面,采用自报和统一抽取相结合,便于构建完整的信息系统;5. 精确计算,分等展示,方便高校对照自己不足和提出改进措施;6. 专家咨询,不断改进,便于与国际发展相接轨。^[5]

正是由于集中性统一化评估模式具有这些优点,才让它在中国高教界通行无阻,而其缺陷往往被人们忽略。事实上,与这种决策快、执行效率高相伴而生的缺陷也非常明显:1. 集中性统一化评估,一旦决策错误就会造成全局性重大影响。这种评估模式往往偏向于终结性评估而非诊断性评估,因为终结性评估易于操作,而诊断性评估难以操作。2. 指标简单统一,无法照顾个性发展需求,容易造成趋同化和恶性竞争。由于指标简单统一,客观地引导人们走向同一模式,按照统一标准来对照先进和发现不足。而简单化的指标最终倾向于由外部发展条件决定,于是学科评估变成了发展条件的相互攀比,从而就造成了恶性竞争。3. 统一评估便于排名,虽然调动了人们的参与热忱,但却使人们把注意力集中到名次上而非学科建设本身,因为名次代表了政绩,而内在建设难以显示政绩,要提高政绩最简捷的办法就是高薪挖角。4. 评估加大了区域间不平衡,加剧了马太效应。通过评估,成绩差者要求进行资源补偿,以便追赶优秀者,但往往受制于资源有限;成绩好要求进行奖励,以便巩固优势,从而进一步强化优势。5. 学科评估强化了大学集中资源意识,进而强化加剧了大学行政化。政府通过学科评估更加强了对学科发展的控制,大学通过学科评估强化了对学科的干预能力。从而,集中的统一评估模式彰显了政府在大学发展过程中的主导性。^[6]

在欧美等发达国家,学科评估一般采用分散的个性化方式进行,具体而言是进行同行评估,由专业协会负责,对评估结果不进行排名,只是为了发现存在哪些不足,以便找到改进的对策,因此是一种诊断性评估。^[7]相对而言,进行排名则属于一种终结性评估,容易对被评对象的声誉造成伤害。欧美虽然也开展一些半官方的评估^①,但都是委托专业协会进行的,政府不直接进行评估,因为政府具有强权威性,是资源的掌控者,直接评估往往会导致政府对学术产生过度干预。准政府机构也享有很大的权威性,采用集中的统一评估模式就容易形成标准化、机械化、刻板化的发展趋势。而同行评估往往采纳的是一种个性化标准,比较尊重各自办学特色。一般而言,这种评估不进行排

^① 在欧洲如德法两国,大学校长协会在高教评估中扮演着极其重要的角色,他们往往是大学与政府关系的协调人,具有准官方的地位。在英国,大学拨款委员会在大半个世纪中都发挥着大学与政府关系协调人的作用,具有准官方的地位。如今高等教育基金会取代了大学拨款委员会的作用,但仍然具有准官方的地位。在美国,区域认证组织在大学评估中发挥着核心作用,而联邦教育认可机构则负责对这些认证组织的资质进行认证,所以认证组织也具有准官方的地位。

名,所进行的只是一种合格性评估,因为学界普遍坦诚很难形成一个统一标准,也无法判定什么标准是最优的或普适性的,只能达成一些基本的一致意见,形成一些基本共识。显然,这种共识是具有弹性的,很难强求一律,而这种弹性标准却激发了学界进行充分交流,并鼓励大学办出特色和学科具有自己的个性。所以,同行评估绝不是为了树立标杆,而是为了维护基本的学术规范,从而达到质量保障目的。

一旦政府直接从事评估或指派部门进行评估,那么评估性质就发生了变化,这种评估往往具有选拔性或评优性,其目的就在于确立某种权威,采用统一标准的意义就在于此。一旦采用统一化标准,那么学校特色和学科个性将失去了意义,这种评估就容易变成一种机械比对,其目标也不在于发现真正问题,其结果是导向同质化、标准化。正因为如此,学术界对统一化评估是内心抵触的,仿佛评估就是为了排名,就是为个性化发展套上紧箍咒。所谓“尺有所短寸有所长”,排名的结果就是让大家各爆其短,使自我声誉受损。人们尤其不喜欢那些不分类的综合性评估,仿佛所有高校都是一个模型的。从理论上说,不同类型高校办学定位是不同的,学科发展标准也应该不同。人们比较欢迎的是进行有针对性的诊断性评估,因为它能够帮助被评对象发现问题,并提出改进的建议。其比较尊重被评估对象的个性,人们参评是自愿的,而且对参评者信息是保密的。很显然,一旦参评者信息被泄露,大学声望就会遭到质疑,就会压迫参评者按照权威标准进行整改,最终结局必然是大学办学走向标准化和同质化及参与恶性竞争。

所以,评估手段往往决定了评估性质。统一化评估就注定了它的终结性,评估价值在于发挥奖惩作用,排名是最直接的体现。^[8]个性化评估决定了它的诊断性,即为了帮助评估对象发现问题,以便找到改进对策。正源于此,诊断性评估往往受到人们推崇。我国学科评估所确立的“以评促建、以评促改、以评促管”宗旨也具有这一价值取向,^[9]无奈采用的是统一化评估模式,从而评估就不属于诊断性而转化为终结性的。学科排名则是一种无形的资源分配体制,显露出明显的终结性。这也是高校出现排名大战的真正原因。由排名大战,引发了盲目攀比、追求大而全,也加剧了大学“千校一面”发展趋势,强化了学校只注重数字提升和不计成本投入及轻视内涵建设的倾向。因为内涵建设很难见效,也不易测量,一旦评估无法促进内涵建设,那么学科评估就是得不偿失。

二、对学科排名科学性与正当性的审问

无疑,学科评估的目的决不是为了排名,但它很容易被异化为排名,因为排名能够激发高校积极参与,因为排名所造成的危机感迫使大家必须高度重视,全力以赴。换言之,如果没有排名,只进行个性化评价,那么学校是否参与就全凭自愿了,此时大学才是自主的。如果大学发展是真正自主的,就会为追求自己内在的卓越而努力,在参与评估时就会表现出高度的选择性,不会参加对自身发展没有实质意义的评估。但一旦进行排名,就会产生胁迫效应,学校就会面临两难选择:如果不参与,就容易被认为害怕评估;而一旦参与,就必须跟着排名走而失去自己的个性。事实上,没有学校不惧怕排名,除非它是巨无霸型的^①,否则总会暴露这样或那样的短处,形成不良后果。大学之所以追求大而全,就是试图从根本上对抗差评带来的后果,降低对学校声望产生的威胁。如果大家都追求大而全,那么,所比拼的就是物质投入的大小,这也是各大学拼命“砸钱”的起因。这也是排名引起的连锁效应。

^① 无论大学排名还是学科排名,总有一些巨无霸型高校,如在国际大学排行榜中,不可能把哈佛、斯坦福或牛津、剑桥排得太差,否则人们会质疑排行榜存在问题而非大学办学问题。在国内大学排名中,北大、清华永远都在前三甲之内,否则就会认为排行榜出了问题。就教育学科而言,如果北师大、华东师大排名差了,则人们会质疑学科评估指标不合理,诸如此类等等。

一般而言,人们的行为预期就决定了其行为选择。因为学科排名直接关乎大学的核心利益,也即大学的声望,从而直接影响到其生源状况和资源状况,所以,大学管理者对于学科排名都异乎寻常地重视。如此,就必须对统一化学科评估的科学性和正当性进行审问。

首先需要审问的问题是:1. 谁能够制定出一个适用于所有高校的学科评估标准? 2. 学科发展是否存在一个公认的、明确的、固定的标准? 不然,制定评估标准的依据何在? 3. 不同学校的学科发展是否应该处在一个统一的发展水平上? 不然排名的意义何在? 4. 学科发展水平是否由发展条件决定? 如果是,学科发展条件能够做到统一吗? 否则,评估意义又是什么? 5. 学科评估是否想缩小发展的条件差异? 学科评估是否有利于缩小发展条件差异? 可以说,这一系列发问关乎学科评估的逻辑基础问题。

其次要审问的问题是:1. 学科评估结果是否属于学校的核心机密,是否必须公布? 因为学科排名就是对学科发展水平的判定,直接影响到学科声望,关乎学校的核心利益,从而属于学校机密范畴。2. 谁有权力要求高校公布自己的核心机密? 换言之,参与评估是否必须参与排名? 评估与排名能够直接相等吗? 3. 当评估涉及高校的核心利益时,谁真的能够做到完全自愿参与评估? 换言之,哪个学校乐于自曝其短? 4. 即使是大学自愿参加,评估单位是否有权公布评估结果? 这一系列发问,所关系的正是高校的自主权问题,也是大学能否实现法治问题,因为这关系到大学自身权益的保护问题。

再次要审问的问题是:1. 目前所采用的评估标准科学吗? 可信吗? 真的公平合理吗? 2. 如果不能做到公平合理,造成伤害谁来负责? 3. 它会不会引发马太效应? 有没有预防马太效应出现的办法? 4. 如果没有,为什么要加剧大学发展的不平衡态势? 这些质疑,实际上是学科评估权威性的质疑,是对其价值立场的质疑,也是对这种制度安排的公平性的质疑。

再其次要审问的问题是:1. 目前学科评估对学科建设真的有促进作用吗? 如果有,促进作用究竟有多大? 2. 如何来制止学科发展走向大而全、趋同化、资源大战? 换言之,对于出现大而全、趋同化和资源大战的后果,评估单位如何回应? 3. 能否对评估本身展开元评估? 由哪些部门来对评估进行元评估? 这些质问,实际上是对传统的大学治理体系的质疑,换言之,是对大学治理体系现代化的呼唤。

最后要审问的问题是:1. 评估者的绝对权威地位是谁授予的? 2. 谁能保证它是站在完全中立的立场上? 换言之,它是否真正具有学科权威性? 3. 是否有学者参与就代表它是客观的、公正的? 可以说,这些质问是对学科评价体制的质疑,是对传统管理体制有效性的质疑,也是对大学发展如何摆脱路径依赖的呼唤。

正是因为学科排名对参评高校产生巨大的利害关系,所以没有人能够真正对排名采取比较客观的理性态度。从人的虚荣心出发,人们都想争最大、最强、最好,谁都不愿意示弱,为了获得最好名次,当然尽可能地把各种指标做好。这会不会影响日常工作? 很显然,对日常工作影响非常大。事实上,在每个学科内部,人们基本上都知道自己的优势是什么,劣势是什么,该与谁比,不该与谁比,无需外部进行评价。但这些又非自己所能左右,往往是身不由己才参加评估。对于学术人而言,基本上都是保持一种谦逊的态度,不想炫耀自己。但要参加评估,就必须竭力地炫耀自己,掩盖自己的不足,不然的话就很容易被“矮化”,那样就容易使自己的声誉受损,此时谦逊不再是美德了,而变成迂腐或愚蠢的代名词了。所以,排名就直接把学术行为导向商业化。

大学既然要参与评估,就不得不按照评估方的规则来进行,因为评估对象是没有选择权的。他们所能做的,只能是竭力地突出自己的优势,想尽一切办法掩盖自己的短处,基本上没有人愿意自曝其短,暴露短处实在是无奈之举。就此而言,玩弄花招就在情理之中了。故而,大学评估与学风变坏之间具有比较直接的关系。难道他们不想诚实吗? 不! 实在是排名对学校声誉影响太大了,而绞尽脑汁掩盖不足就是一种本能冲动。难道大学不想弥补自己的短处吗? 当然想! 但从自

己的发展条件和发展定位而言似乎又很难弥补。这样就造成了两难选择:如果追求大而全就失去了自己的特色,使自己变得平庸;如果不追求大而全,就使自己名誉受损。

在这一点上,国外大学的评估相对更自主一些,因为他们可以根据自己的情况选择究竟评估什么。国内评估采用的是“绑定原则”,^[10]即要求参与者按照学科大类要求,要么一级学科全部参与,要么不参与,从而学校很难抉择:不参与意味着胆怯,似乎有碍自尊;参与则面临风险,可能使名誉受损。这也导致了一种异化现象:在国内,参评的目的是展示,而非发现问题;在国外,参评目的主要是发现问题以便于获得更大发展。这样就出现了一个怪象:知道自己强的学科才愿意参加评估,以便于展示自己;而知道自己存在不足的学科往往不愿意参加评估,生怕参评是自取其辱。

即使那些有底气的学科,为了参评,也必须进行全方位的精心准备,所展示的均非常态,其出发点无非是为了展示最精彩的自己,从另一个角度而言也是为了避免排名伤害的最大化。如此就出现了一系列迎评高手,他们在各处传递经验,而聘请者如云,目标是实现评估利益最大化和评估损害最小化。由此也兴起了一个迎评产业。

人们一般都承认,学科表现如何很大程度上是由发展条件决定的,虽然与自身努力分不开,但环境的决定性是无法否认的,而学科评估并不面对这些问题。如果学科发展条件不平衡,而评估采取的是统一政策,那么这种评价公平性就值得质疑。因为进行公平合理的评估必须确立它的前提:1. 各个学科点发展条件是基本相同的;2. 各个学科发展过程是可以控制的;3. 各个学科点发展目标是一样的,唯有如此,才能采用统一的评估办法,否则,评估是违反伦理的,没有意义的。我们知道,各个学科点的发展条件是不可能相同的,也可以说是差异悬殊的;学科发展过程也是不受控制的,因为人们并不真正掌握学科发展规律,只是了解一些相关因素而已。各个学科点发展目标也不应该是相同的,应该体现出办学层次的差别和办学方向的差别。所有这些问题是各个学科点在发展中必然遇到的。

既然如此,学科评估的价值就导向必须进行审问。^[11]显然,如果为了了解学科发展的一般状况,就无需如此大动干戈,因为这些数据未必真实。如果是为了分等,自然也不必,因为差距是天然存在的。是为了重新分配资源,实现均衡化发展?恐怕这并非评估的目的所在。实质上,评估的直接后果就是符号资源再分配,所体现的仍然是一种马太效应。

三、学科评估影响学科发展的机理

学科评估的真正目的是什么?人们惯常用“以评促建、以评促改和以评促管”来描述。从现实情况看,我国目前学科评估确实引起高校管理层的高度重视,纷纷根据学科评估结果找差距,想对策,这一状况确实是可喜的一面。^[12]但冷静分析后不难发现,各个高校的学科建设对策无外乎增加投入、吸引人才和制定绩效考核政策,目标都在于提高学科排名位次,至于学科发展的内在机制是什么,如何尊重学科特性,如何对学科进行科学评估,这些问题反倒很少人提及,原因就在于对这些问题很难形成共识,也很难采取什么对策。

显然,如果缺乏科学的先进的理念指导,学科建设基本上是盲目的。事实上,大多数高校对于自身学科发展存在的短板是比较清楚的,而且也知道这种短板形成过程,但却很难改变,“非不愿也,实不能也”,这代表了绝大多数高校的心声。因为学科发展问题牵涉复杂的利益关系问题,面对复杂的利益纠葛,学校管理层也往往有心无力。所以,要改变高校内部的利益关系布局,很大程度上取决于学校领导人的魄力和决心。一般而言,学科带头人选不好,则学科建设成效就无法期待。但这些都牵涉深度的院系政治考量,即使学校管理者也不希望深度介入。如此,要打破学科发展惯性谈何容易?

故而,“以评促建、以评促改、以评促管”往往很难进入深层。要使学科评估发挥深层的促进作

用,必须从引导学科发展走向科学发展做起,从培养优良的学术生态的根底做起,也即从促进高校之间开展良性竞争做起。如果不能培养大学的自主发展意识,不能尊重学科发展的特殊规律,不能尊重学术自由的价值,那么学科评估效力是有限的。

如前所言,学科评估面临的最大问题是伦理问题,即评估机构是否有权公布参评单位的信息。如果没有参评单位的授权,又不可避免结果公布给高校发展造成的伤害,那么公布行为就是不当的。因为一旦公布评估结果,必然对高校学术生态产生连锁反应,如同质化和攀比效应。所以,参评高校普遍都处于一种两难处境:如果不参评,则会被认为惧怕评估;而一旦参评,就必然陷入经费大战、人才大战和关系大战之中。因为无论怎么评,总能够发现自己的不足,就要努力改进,无论这种不足是否属于正常范围之内。既然要改进,那么就必须增加投入,无论人、财、物,还是打通各种人脉关系,这就是恶性竞争的开始。我们知道,“金无足赤人无完人”,学科评估似乎是为了打造一种完美学科,这显然是荒诞的。而一旦参与评估,学科发展就丧失了自主性,就只能按照统一指标去做,而且只能假定这些指标设计都是正确的,这就产生了所谓的“对标”效应。这种对标的直接后果就是趋同化、同质化。

无论哪一个学科都存在自己的不足,都有自身的问题,这些问题往往是深层的,唯有自己才最清楚,而评估数据却无法显示这些,也无法为此提供解答答案。很多时候,这种不足也成为一种隐私,不希望被外界窥探。一旦参评的话,这些隐私和痛处势必被触及,甚至成为矛盾的导火索。学科特色从根本上讲都是由特殊的历史经历塑造的,学科评估往往是对这种历史进行审视,从本质上讲就是在解构学科文化,进而引导学科走向技术化和统一化,故而是一种去文化的策略。所以,从伦理性角度考虑,学科评估不应发挥“去文化”效应,仍应采取尊重个性的策略。而集中性的统一化评估显然很难做到这一点,只有进行个别化的诊断性评估才能满足这一要求。

尽管如此,集中性统一化评估仍有一定价值,其价值就在于在学科发展初期阶段提供一个基本的学科规范。因为许多高校在学科建设上缺乏明确方向,统一化评估能够为他们指明规范化方向。不少新建本科学校,由于缺乏学术积淀,对学科建设不了解,进行统一化学科评估有助于学科规范的普及。由此可见,统一化评估仅适用于合格性评估,不适合于选优性评估,因为真正优秀都是个性化非常强的,是不可能走统一路线的。

我们必须承认,统一化的学科评估在促进高校学科规范化建设中能够产生外部督促效应。现实中发现,不少高校在学科建设中容易陷于利益分配的漩涡中,外部评估能对这些行为产生一定的抑制作用。毋庸置疑,统一化学科评估好像是对学科发展基本状况进行一次系统性的体检,从中发现相对优势和不足。一旦进行排名,这些优缺点都容易被放大,从而引起高校领导人的高度重视。正是学科排名所具有的广告效应,促使人们疯狂地追逐排行榜。然而,由于它无力为学科发展状况做出科学诊断并提出针对性建议,其促进作用是相当有限的,而其负作用却是非常突出的,即引导高校盲目攀比,走趋同化发展路线。

一般认为,学科评估的真正价值恰恰在于能够对学科发展状况做出科学的诊断。显然,要做出科学的诊断,就需要建立符合学科特征的评估指标体系,而且需要掌握学科发展的全方位的准确信息,特别是要了解学科发展的历史及其特性。如果参评对象不主动提供真实信息的话,就不可能做出科学的诊断。“以评促建、以评促管、以评促改”的前提假设是它能够进行科学的诊断,而且也知道该如何改进,但这个假设是无法证实的。因为学科发展环境是复杂的,并无统一模式,很难做出科学诊断,也无法找到问题的真正根源,当然也无从进行科学的指导。学科自身也是一个生命体,它有自身的独特价值追求,而且在不同环境下也会产生出不同的个性,如何尊重学科自身的生命价值是学科评估必须思考的问题。

无法否认,管理部门对学科发展状况具有测评权力,^[13]以便制定学科发展规划和确定学科发展战略,但必须在保证促进学科发展的前提下进行。如果仅仅为了排名,无法提供科学的咨询服

务,不能鼓励办学走向个性化和特色化,则评估意义就值得怀疑。^①即使进行排名,也必须做到科学、公正、合理,首先是要进行投入状况与建设成效进行比对,否则评估就是一笔糊涂账。显然,要对投入与产出效益进行评估,绝非管理部门自身所能够完成的,而且学术投入与产出关系往往是不可预期的,这正是评估的难点所在。

必须承认,目前集中性的统一化学科评估具有明显的局限性。就像目前的各种医疗设备一样,虽然它能够监测出一定的问题,但往往难以给出合理的解释,而解释水平又往往取决于医生的临床经验。学科评估也是如此,它只能告诉人们某些指标高低,但无法说明这种高低是否属于合理范围内,也不能揭示其必然原因。所以,对于目前的学科评估热,需要冷静看待,不能一切跟着评估指标走。^[14]

四、第四轮学科评估中的问题反思

无论是学科负责人还是学科发展规划设计者,都清楚学科发展在许多方面是不可控制的,而且学科发展条件也基本上是不可比的,因为学科发展不可避免地受学校所在的地域影响和领导人的思维模式影响。换言之,所在地区的社会经济发展水平和高校领导人的思维方式基本上就决定了高校发展状态和学科发展水平,而学科评估无法解决这些问题。我们不否认学科评估可以起到一定的督促作用,但其负作用也影响很大,最典型的就是引导高校趋同化发展并助长了高校外延式发展惯习,同时也加剧了人们的浮躁心理,让人们投入巨大精力去做排名,而不是扎扎实实做好学科建设。^[15]此外,学科评估也强化了行政干预能力,使大学行政化趋势更加积重难返,因为大学要提升学科排名,就必须加强控制,就更不鼓励自由探索了。在急于求成的心态下,学科发展逻辑和学术基本规律就被置于脑后。当行政部门把层层指标强加在教师们头上时,大学实际上就违背了学术发展基本规律。

就第四轮学科评估的实际影响而言,它给人的感觉是五味杂陈。^[16]首先,与之前的学科评估相比,它确实有不少进步,如去掉了许多“帽子人才”所占的分值,这对于抑制各个高校的抢人大战是有益的,体现了一种正确导向。^[17]不过,这很难产生一种治本效果,因为它无法从根本上抑制高校进行人才储蓄的欲望,原因在于各种帽子人才背后潜藏着巨大的资源,其诱惑力是巨大的。所以,“去掉帽子人才”无法改变大学之间不平等竞争的事实,而且从长远看,“帽子人才”会比以往受到更大的重视。

其次,采用绑定原则可谓利弊难料。不可否认,绑定原则的出发点无疑是好的,因为它规定如果同一门类的的一个一级学科参评,其他一级学科都必须参评。如此可避免同一成果被重复使用,从而提高了评估的可信性。^[18]但绑定原则所产生的效应是:规模越大则优势越大,如此就把学科建设导向大而全,因为只有达到足够规模,才能在各项指标中占据有利地位。显然,师资规模越大,所产生的论文、专著和获奖数以及课题立项数就会越多,因为它不评比“人均贡献”,从而鼓励学科发展重全而不重精。只有进行人均产出评比时,学科建设才能走向小而精,而这是走向个性化、特色化的必经之路。不知为何,如此重要的指标被忽略了,显然这是评估导向出现了差错。

再次,评估结果出现了一系列令人费解的现象。如有的学科声誉很高,却培养过程质量很低,但毕业生很优秀,这个矛盾现象真让人百思不得其解;有的学科生源质量低而毕业生质量却很高,

^① 对于评估实施主体而言,他们往往并不拥有学术权威,所以他们对学科评估指标体系的确定不得不借重学科专家的意见。但他们不可能征求全部专家的意见,只能抽取其中的极少一部分人的意见,这部分人如何选择就大有考究。一般而言,人们不可能按照严格科学原则抽取专家,更倾向于方便抽样,如此专家意见就大打折扣。对于专家们的意见,评估实施主体未必尽能领会,这样的话对于评估结果也不可能进行非常有效的解释。特别是实施主体也无法长期聘请专家作为评估结果的解释者并对参评者开展咨询服务,如此“以评促建”愿望势必落空。

这只能说明培养过程质量很高,但培养过程质量得分却很低!显然,这要么是评估指标设计出了问题,要么是评估手段不科学造成的。对于这种反常现象,评估部门往往无法给以解释。最让人感到吃惊的是,没有本科生的学科普遍出现了“培养过程质量低谷”现象,而有本科生教育的评估得分普遍较高,这似乎说明学科评估是针对本科生教育而非研究生教育,评估导向是为了鼓励所有学科必须举办本科教育。

这一系列反常现象说明了一个问题:评估指标设计不科学,没有进行元评估。^[19]要知道,学科评估主要是针对研究生教育的,而非针对本科生教育的。事实证明,没有本科生教育的学科,在精品课程、精品教材和网络课程的成绩普遍落后。而研究生教育并不适于搞精品课程和精品教材及网络课程,因为研究生教育目的并非为了传授现成知识,而是为了培养创新能力。一般而言,教材是知识定型化的结果,也是进行系统传授的依据,研究生教育是以能力培养为主,而精品课程与精品教材容易导致标准化教学和灌输式教学,并不利于创新人才培养。如果指标设计落后于时代发展要求,则会产生明显的误导作用。

鉴于此,学科评估必须站在研究生教育发展前沿,以鼓励创新人才作为自己的使命,不应引导学科建设走向标准化和统一化。

五、结论:不同学科应采用不同评估模式

学科评估对于学科发展意义非同寻常,无论采用集中性统一化评估,还是采用分散性个别化评估,都必须根据学科发展需要而定。集中性统一化评估,由于标准统一、时间集中、结果排名,从而有利于学科开展竞赛,有利于学科建设的规范化,但其弊处在于无法针对学校的个性需求,无法鼓励学校的特色发展。而个性与特色乃是大学和学科发展的生命,因为这是创造力的源泉。分散性个别化评估,操作的难度大,效率低,成本高,比较适合于诊断性评估,有利于大学的个性和学科的特色发展。对于大学而言,没有个性就会变得平庸;对于学科而言,没有特色就失去了存在价值。

一般而言,学科内部的共识性越强,越适于采用集中性统一化评估模式;相反,如果学科内部缺乏共识,则适于采用分散性个别化评估。学科内部之所以能够形成共识,就在于学科发展范式比较成熟,如有公认的概念体系和研究方法,也有深厚的学科文化积累。这类学科一般而言属于硬科学,如物理学。相反,如果学科没有形成公认的研究范式,在基本概念上缺乏一致意见,在研究方法上是多元的甚至是相互冲突的,那么这样的学科属于软科学,人文科学和多数社会科学大多如此。这说明,对于自然科学而言,统一化的学科评估有利于找到成功的经验和反思自身不足,从而有利于确立改进的方向和目标,最终促进学科整体进步,因而利大于弊。而对于人文科学和大多数社会科学而言,统一化评估往往会导致学科发展走向僵化,其弊处远大于利处。

我们可否设想对各个学科进行特色排名?如果能够对学科特色进行排名,就能够大大消除人们对排名的恐惧,就能够极大地鼓励各个学科走向创新,这样就能够极大地激发各个高校的办学活力。因为没有特色就不可能成为一流,而有无特色尚需学界公认。由学界同行对学科特色进行评价,无疑会激发各个高校的自主性和创造性。如此学科评价才能真正发挥促进“双一流”建设的效应,就会避免走入趋同化、行政化的窠臼。

注释:

[1]王洪才《“双一流”建设的重心在学科》,《重庆高教研究》2016年第1期。

[2]马陆亭《一流学科建设的逻辑思考》,《高等工程教育研究》2017年第1期。

[3][4][10][13][17][18]王小梅等《以学科评估为契机 提升学科建设水平》,《中国高教研究》2016年第12期。

- [5]王立生等《我国学科评估的发展历程和改革探究》,《中国高等教育》2016 年第 21 期。
- [6]陈学飞等《中国式学科评估: 问题与出路》,《探索与争鸣》2016 年第 9 期。
- [7]蒋林浩等《学科评估的方法、指标体系及其政策影响: 美英中三国的比较研究》,《高等教育研究》2014 年第 11 期。
- [8]林梦泉等《学科评估发展与改革探究》,《中国高等教育》2010 年第 21 期。
- [9]黄宝印《努力构建中国特色国际影响的学科评估体系》,《中国高等教育》2018 年第 1 期。
- [11]熊庆年、张瑞鸿《学科评估中的价值评估及其类型》,《教育发展研究》2017 年第 3 期。
- [12]袁本涛、李锋亮《对我国学科评估发展的调查与分析》,《高等教育研究》2016 年第 3 期。
- [14]王建华《一流学科评估的理论探讨》,《大学教育科学》2012 年第 3 期。
- [15]翟亚军、王晴《“双一流”建设语境下的学科评估再造》,《清华大学教育研究》2017 年第 6 期。
- [16]刘尧《“双一流”建设评估困境何以突破——从全国第四轮学科评估结果引起舆论风波谈起》,《江汉大学学报(社会科学版)》2018 年第 2 期。
- [19]赵立莹《问责与改进: 我国第四轮学科评估元评估》,《学位与研究生教育》2018 年第 2 期。

[责任编辑: 蔡永明]

Discipline Rankings: Help or Hinderance?

——Reflections on the Feature, Fairness and Mechanism of China's Discipline Assessment

WANG Hong-cai

(Center of Higher Education Development Research, Xiamen University, Xiamen, 361005, Fujian)

Abstract: In China, discipline assessment is a priority for universities and colleges because of its intimate relationship with the “Double World-class” building. Undoubtedly, China's discipline assessment is administratively organized and is therefore centralized and unified, and as a basic tool, discipline rankings are highly attractive to universities, but the practice compels universities to irrationally compare with one another and leads to homogeneity. Although it has made a technical improvement in the Fourth Discipline Assessment, the new “binding principle” is at odds with the original intention of “voluntary participation”, the principle of “first-level discipline” assessment is disadvantageous to the intensive development of disciplines. Therefore, Chinese disciplines need a series of reexaminations on feature, fairness and mechanism. This paper concludes that discipline assessment should be diagnosis-oriented, pose no threat to university autonomy, and respect the individuality of universities and the characteristics of disciplines, in order to facilitate attainment of world-class status. Otherwise, the assessment will do more harm than good.

Keywords: discipline assessment, discipline rankings, “Double World-class” building